The theory of minimax approximation

Huaijin Wang

September 11, 2024

Contents

1 Introduction to minimax approximation **2**

2 The reduction of the error of a trial approximation **3**

3 The characterization theorem and the Haar condition **7**

4 Uniqueness and bounds on the minimax error **13**

5 Appendix **16**

1 Introduction to minimax approximation

Let $\mathscr{C}[a,b]$ be the set of all continuous functions that are defined on the interval $[a,b]$ and $\mathscr{A}$ is a linear subspace of $\mathscr{C}[a,b]$ with finite dimensional.

The best minimax approximation from $\mathscr{A}$ to a function f in $\mathscr{C}[a,b]$ is the element of $\mathscr{A}$ that minimizes the expression

$$\|f-p\|_\infty = \max_{a\leqslant x\leqslant b} |f(x)-p(x)|, \quad p\in\mathscr{A}. \tag{1}$$

Thus, the best minimax approximation problem is given a function $f \in \mathscr{C}[a,b]$,

$$\begin{cases} \text{Find } p^* \in \mathscr{A} \text{ s.t.} \\ \|f-p^*\|_\infty \leqslant \|f-p\|_\infty, \quad \forall p\in\mathscr{A}. \end{cases} \tag{2}$$

In this notes we study the conditions that are satisfied by a best approximation, when $\mathscr{A}$ is a general linear space.

We note that they take a particularly simple form if $\mathscr{A}$ is the space $\mathscr{P}_n$ of algebraic polynomials of degree at most n.

In fact this form is obtained in the more general case when $\mathscr{A}$ satisfies the 'Haar condition', which is defined in Section 7.3. In Section 7.4 some further useful properties of best minimax approximations are proved in the case when the Haar condition is obtained, including the result that the best approximation is unique.

The Haar condition also provides an excellent method for calculating best approximations, called the exchange algorithm, which is described in Chapter 8 and analysed in Chapter 9.

The theory that is developed for the case when $\mathscr{A}$ is any finite dimensional linear space comes from asking the following question. Let p^* be a trial approximation from $\mathscr{A}$ to f. Can we find a change to p^* that reduces the maximum error of the trial approximation? In other words, we seek an element p in $\mathscr{A}$ such that the inequality

$$\|f-(p^*+\theta p)\|_\infty < \|f-p^*\|_\infty \tag{3}$$

is satisfied for some value of the scalar parameter θ. Figure 7.1 gives an example to explain this point of view.

In the figure the function f, which is shown in each of the four parts, is to be approximated by a straight line, so $\mathscr{A}$ is the space $\mathscr{P}_1$. Three trial approximations, namely p_1^*, p_2^* and p_3^*, are shown. The vertical lines in the figure indicate where the error function of each approximation takes its

maximum value. We see that the straight line p_1^* is not optimal, because the maximum error is reduced if the line is raised. The straight line p_2^* is not optimal either, because the maximum error can be reduced

by rotating the line in a counter-clockwise direction. The straight line p_3^*, however, is the best approximation from $\mathscr{P}_1$ to f. We find in Section 7.3 that the characteristic property of a best straight line approximation is that the maximum error is achieved at three points of $[a, b]$ with alternating sign.

Figure 7.1 suggests that, to discover if a trial approximation is optimal, one only need consider the extreme values of the error function $\{f(x) - p^*(x); a \leqslant x \leqslant b\}$. This remark is made rigorous in the next section. It follows that we can find a function, g say, to add to the function of Figure 7.1, such that the best approximation is unchanged, but the best approximation from $\mathscr{P}_1$ to g is not the zero function. This remark is important, because it shows that in general a best minimax operator from $\mathscr{C}[a, b]$ to $\mathscr{A}$ is not a linear operator. Therefore the algorithms for calculating best approximations are iterative procedures.

2 The reduction of the error of a trial approximation

In this section, we consider the problem of how to deduce a "trial approximation" to a "better approximation".

Let p^* be a trial approximation from $\mathscr{A}$ to a function f in $\mathscr{C}[a, b]$, and we try to improve the approximation by satisfying condition (3). The set of points at which the error function

$$e^*(x) = f(x) - p^*(x), \quad a \leqslant x \leqslant b, \tag{4}$$

takes its extreme values is important, and we call it $\mathscr{Z}_M$, i.e.,

$$\mathscr{Z}_M = \{x \in [a, b] : |e^*(x)| = \|e^*\|_\infty\}. \tag{5}$$

We suppose first that p^* is not optimal. We let $(p^* + \theta p)$ be a best approximation. Hence the reduction (3) is obtained, and the points in $\mathscr{Z}_\mathrm{M}$ satisfy the inequality

$$|e^*(x) - \theta p(x)| < |e^*(x)|, \quad x \in \mathscr{Z}_\mathrm{M}. \tag{6}$$

We assume without loss of generality that θ is positive. Therefore expression (7.5) shows that, if x is in $\mathscr{Z}_\mathrm{M}$, then the sign of $e^*(x)$ is the same as the sign of $p(x)$. In fact, this result is simplified as the conclusion: for any $a, b \in \mathbb{R}$, then $|a - b| < a$ implies that the sign of a is the same as the sign of b.

Lemma 2.1. *p^* is a best minimax approximation from $\mathscr{A}$ to f if there is no function p in $\mathscr{A}$ that satisfies the condition*

$$[f(x) - p^*(x)]\, p(x) > 0, \quad x \in \mathscr{Z}_{\mathrm{M}} \tag{7}$$

Proof. If there is no function $p \in \mathscr{A}$ s.t. (7) holds, namely $\forall p \in \mathscr{A}$, we have

$$e^*(x)p(x) \leqslant 0, \quad \forall x \in \mathscr{Z}_M,$$

where $e^*(x)$ is defined as in (4). Then

$$\mathrm{sign}(e^*(x)) = -\mathrm{sign}(p(x)), \quad \forall x \in \mathscr{Z}_M.$$

It leads to $\forall p \in \mathscr{A}$ and $\forall \theta > 0$,

$$|e^*(x) - \theta p(x)| \geqslant |e^*(x)|, \quad \forall x \in \mathscr{Z}_M. \tag{8}$$

Replace $(p^* - \theta p)$ with p and take the maximum over $\mathscr{Z}_M$ on both sides of (8), leading to

$$\max_{x \in \mathscr{Z}_M} |f(x) - p^*(x)| \leqslant \max_{x \in \mathscr{Z}_M} |f(x) - p(x)|, \quad \forall p \in \mathscr{A}.$$

Therefore, $\forall p \in \mathscr{A}$,

$$\begin{aligned} \max_{a \leqslant x \leqslant b} |f(x) - p^*(x)| &= \max_{x \in \mathscr{Z}_M} |f(x) - p^*(x)| \\ &\leqslant \max_{x \in \mathscr{Z}_M} |f(x) - p(x)| \leqslant \max_{a \leqslant x \leqslant b} |f(x) - p(x)|. \end{aligned}$$

□

In the remainder of this section the converse result of Lemma 2.1 is proved, namely that, if inequality (7) holds for some p in $\mathscr{A}$, then there exists a positive value of θ that gives the reduction (3).

Because of the way in which the exchange algorithm works, we generalize the problem of minimizing $\|f - p\|_\infty$, to be a more general case.

Let $\mathscr{Z}$ be any closed subset of $[a, b]$, which may be $[a, b]$ itself, or just composed of a finite number of points. Then, we consider the problem: for given $f \in \mathscr{C}[a, b]$,

$$\begin{cases} \text{Find } p^* \in \mathscr{A} \text{ s.t.} \\ \max\limits_{x \in \mathscr{Z}} |f(x) - p^*(x)| \leqslant \max\limits_{x \in \mathscr{Z}} |f(x) - p(x)|, \quad \forall p \in \mathscr{A}. \end{cases} \tag{9}$$

Remark 2.1. *Compared with the problem of* (2)*, the problem of* (9) *would be more general and focus on the the local property which will be more suitable for computation.*

Theorem 1. *Let $\mathscr{A}$ be a linear subspace of $\mathscr{C}[a,b]$, let f be any function in $\mathscr{C}[a,b]$, let $\mathscr{Z}$ be any closed subset of $[a,b]$, let p^* be any element of $\mathscr{A}$, and let $\mathscr{Z}_{\mathrm{M}}$ be the set of points of $\mathscr{Z}$ at which the error $\{|f(x)-p^*(x)|\,;x\in\mathscr{Z}\}$ takes its maximum value, i.e.,*

$$\mathscr{Z}_M=\{x^*\in\mathscr{Z}:|f(x^*)-p^*(x^*)|=\max_{x\in\mathscr{Z}}|f(x)-p^*(x)|\}.$$

Then $p^*\in\mathscr{A}$ is a minimax approximation that satisfies (9) *if and only if there is no function p in $\mathscr{A}$ that satisfies condition* (7)*.*

Proof. **Sufficiency.** The proof is really the same as it does in Lemma 2.1. Suppose that there is no function $p\in\mathscr{A}$ that satisfies (7), then $\forall p\in\mathscr{A}$ there holds

$$[f(x)-p^*(x)]p(x)\leqslant 0,\quad \forall x\in\mathscr{Z}_M.$$

Namely, the sign of $p(x)$ is different to the sign of $e^*(x)$ as x ranges over $\mathscr{Z}_M$, i.e.,

$$\mathrm{sign}(e^*(x))=-\mathrm{sign}(p(x)),\quad \forall x\in\mathscr{Z}_M.$$

It leads to $\forall p\in\mathscr{A}$ and $\forall\theta>0$,

$$|e^*(x)-\theta p(x)|\geqslant|e^*(x)|,\quad \forall x\in\mathscr{Z}_M.$$

Then, $\forall p\in\mathscr{A}$,

$$\max_{x\in\mathscr{Z}_M}|f(x)-p(x)|\geqslant\max_{x\in\mathscr{Z}_M}|f(x)-p^*(x)|.$$

Hence $\forall p\in\mathscr{A}$,

$$\begin{aligned}\max_{x\in\mathscr{Z}}|f(x)-p^*(x)|&=\max_{x\in\mathscr{Z}_m}|f(x)-p^*(x)|\\&\leqslant\max_{x\in\mathscr{Z}_M}|f(x)-p(x)|\leqslant\max_{x\in\mathscr{Z}}|f(x)-p(x)|.\end{aligned}$$

Necessity. We prove it by contraction. Suppose that $p\in\mathscr{A}$ that satisfies (7). We need to show that there exists $\theta>0$ such that

$$\max_{x\in\mathscr{Z}}|e^*(x)-\theta p(x)|<\max_{x\in\mathscr{Z}}|e^*(x)|. \tag{10}$$

Without loss of generality, we assume

$$|p(x)|\leqslant 1,\quad \forall x\in[a,b].$$

Let

$$\mathscr{Z}_0 := \{x_0 \in \mathscr{Z} : p(x_0)e^*(x_0) \leqslant 0\}.$$

It is obvious that $\mathscr{Z}_0$ is closed, and because $\mathscr{Z}_0$ and $\mathscr{Z}_M$ have no common points, the number

$$d = \max_{x \in \mathscr{Z}_0} |e^*(x)|$$

satisfies the bound

$$d < \max_{x \in \mathscr{Z}} |e^*(x)|.$$

If $\mathscr{Z}_0 = \varnothing$, we denote $d = 0$. (Note that the inequality strictly holds since $|e^*(x)|$ achieves its maximum on $\mathscr{Z}_M$, but not $\mathscr{Z}_0$, and $\mathscr{Z}_0 \cap \mathscr{Z}_M = \varnothing$).

We will show that (10) holds with θ as

$$\theta = \frac{1}{2}\left[\max_{x \in \mathscr{Z}} |e^*(x)| - d\right].$$

Because $\mathscr{Z}$ is closed, there must be a $\xi \in \mathscr{Z}$ such that

$$|e^*(\xi) - \theta p(\xi)| = \max_{x \in \mathscr{Z}} |e^*(x) - \theta p(x)|.$$

- If $\xi \in \mathscr{Z}_0$, the bound

$$\begin{aligned}\max_{x \in \mathscr{Z}} |e^*(x) - \theta p(x)| &= |e^*(\xi)| + |\theta p(\xi)| \leqslant d + \theta \\ &= \frac{1}{2}\left[\max_{x \in \mathscr{Z}} |e^*(x)| + \max_{x \in \mathscr{Z}_0} |e^*(x)|\right] \\ &< \max_{x \in \mathscr{Z}} |e^*(x)|,\end{aligned}$$

  where the first equal holds because the sign of $e^*(x)$ is the same as the sign of $p(x)$ as we assume that $p(x)$ satisfies (7).

- If $\xi \notin \mathscr{Z}_0$, then $\operatorname{sign}(e^*(\xi)) = \operatorname{sign}(p(\xi))$ and

$$|e^*(\xi) - \theta p(\xi)| < \max\left[|e^*(x)|, |\theta p(\xi)|\right].$$

  By the facts that

$$\begin{aligned}&|e^*(\xi)| \leqslant \max_{x \in \mathscr{Z}} |e^*(x)|, \\ &|\theta p(\xi)| \leqslant \theta = \frac{1}{2}\left[\max_{x \in \mathscr{Z}} |e^*(x)| - \max_{x \in \mathscr{Z}_0} |e^*(x)|\right] \leqslant \max_{x \in \mathscr{Z}} |e^*(x)|.\end{aligned}$$

  It follows that (10) is satisfied.

$\square$

Theorem 1 suggests that to find out if a trial approximation is optimal, one only need consider the extreme values of the error function. Specifically, one should ask if condition (7) holds for some function p in $\mathscr{A}$.

3 The characterization theorem and the Haar condition

If the set $\mathscr{A}$ of approximating functions is the space $\mathscr{P}_n$ of algebraic polynomials of degree at most n, then it is rather easy to test whether condition (7) can be obtained. We make use of the fact that a function in $\mathscr{P}_n$ has at most n sign changes. Therefore, if the error function $[f(x) - p^*(x)]$ changes sign more than n times as x ranges over $\mathscr{Z}_{\mathrm{M}}$, then p^* is a best approximation. Conversely, if the number of sign changes does not exceed n, then we can choose the zeros of a polynomial in $\mathscr{P}_n$ so that condition (7) is satisfied. This result is usually called the minimax characterization theorem, and it is stated formally below.

It is useful to express the theorem in a form that applies to a class of functions that includes polynomials as a special case. The usual way of defining this class is to identify the properties of polynomials that are used in the proof of the characterization theorem. They are the following two conditions:

(1) If an element of $\mathscr{P}_n$ has more than n zeros, then it is identically zero.

(2) Let $\{\zeta_i; j = 1, 2, \ldots, k\}$ be any set of distinct points in the open interval $(\ a, b\ )$, where $k \leqslant n$. There exists an element of $\mathscr{P}_n$ that changes sign at these points, and that has no other zeros. Moreover, there is a function in $\mathscr{P}_n$ that has no zeros in $[a, b]$.

The following two properties of polynomials are required later:

(3) If a function in $\mathscr{P}_n$, that is not identically zero, has j zeros, and if k of these zeros are interior points of $[a, b]$ at which the function does not change sign, then the number $(j + k)$ is not greater than n.

(4) Let $\{\xi_j; j = 0, 1, \ldots, n\}$ be any set of distinct points in $[a, b]$, and let $\{\phi_i; i = 0, 1, \ldots, n\}$ be any basis of $\mathscr{P}_n$. Then the $(n + 1)\times\ (n + 1)$ matrix whose elements have the values $\{\phi_i\,(\xi_j)\,;\ i = 0, 1, \ldots, n; j = 0, 1, \ldots, n\}$ is non-singular.

Remark 3.1. *Condition (2) states that there exists two kind of polynomials in $\mathscr{P}_n$ such that only k roots at which sign changes contained in (a, b), or no zeros in $[a, b]$.*

An $(n+1)$-dimensional linear subspace $\mathscr{A}$ of $\mathscr{C}[a, b]$ is said to satisfy the *'Haar condition'* if these four statements remain true when $\mathscr{P}_n$ is replaced by the set $\mathscr{A}$. Let's rewrite the definition just to be a little bit more specific:

Definition 3.1 (Haar Condition)**.** *The $(n+1)$-dimensional linear subspace $\mathscr{A}$ of $\mathscr{C}[a, b]$ is said to satisfy the Haar condition if these four statements remain true*

(1) If an element of $\mathscr{A}$ has more than n zeros, then it is identically zero.

(2) Let $\{\zeta_i; j = 1, 2, \ldots, k\}$ be any set of distinct points in the open interval (a, b), where $k \leqslant n$. There exists an element of $\mathscr{A}$ that changes sign at these points, and that has no other zeros. Moreover, there is a function in $\mathscr{A}$ that has no zeros in $[a, b]$.

(3) If a function in $\mathscr{A}$, that is not identically zero, has j zeros, and if k of these zeros are interior points of $[a, b]$ at which the function does not change sign, then the number $(j + k)$ is not greater than n.

(4) Let $\{\xi_j; j = 0, 1, \ldots, n\}$ be any set of distinct points in $[a, b]$, and let $\{\phi_i; i = 0, 1, \ldots, n\}$ be any basis of $\mathscr{A}$. Then the $(n+1)\times(n+1)$ matrix whose elements have the values $\{\phi_i(\xi_j); i = 0, 1, \ldots, n; j = 0, 1, \ldots, n\}$ is non-singular.

Remark 3.2. *Equivalently, any basis of $\mathscr{A}$ is called a 'Chebyshev set'. Spaces that satisfy the Haar condition are studied in Appendix A. It is proved that properties (1), (3) and (4) are equivalent, and that these properties imply condition (2). It is usual to define the Haar condition in terms of the first property. Thus $\mathscr{A}$ satisfies the Haar condition if and only if, for every non-zero p in $\mathscr{A}$, the number of roots of the equation $\{p(x) = 0 : a \leqslant x \leqslant b\}$ is less than the dimension of $\mathscr{A}$.*

Remark 3.3. *It is also useful to define the Haar condition in terms of the Chebyshev system: $\mathscr{A}$ satisfies the Haar condition if and only if any basis with $(n+1)$ elements of $\mathscr{A}$ constitutes a Chebyshev system.*

Theorem 2 (Characterization Theorem)**.** *Let $\mathscr{A}$ be an $(n+1)$-dimensional linear subspace of $\mathscr{C}[a, b]$ that satisfies the Haar condition, and let f be any function in $\mathscr{C}[a, b]$. Then p^* is the best minimax approximation from $\mathscr{A}$ to*

f, if and only if there exist $(n+2)$ points $\{\xi_i^* : i = 0, 1, \ldots, n+1\}$, such that the conditions

$$a \leqslant \xi_0^* < \xi_1^* < \ldots < \xi_{n+1}^* \leqslant b, \tag{11}$$

$$|f(\xi_i^*) - p^*(\xi_i^*)| = \|f - p^*\|_\infty, \quad i = 0, 1, \ldots, n+1 \tag{12}$$

and

$$f\left(\xi_{i+1}^*\right) - p^*\left(\xi_{i+1}^*\right) = -\left[f(\xi_i^*) - p^*(\xi_i^*)\right], \quad i = 0, 1, \ldots, n \tag{13}$$

are obtained.

Proof. **Necessity.** Suppose that p^* is the best minimax approximation from $\mathscr{A}$ to f. Let $\mathscr{Z}_M$ be the set for which $|f(x) - p^*(x)|$ attains its maximum, i.e.,

$$\mathscr{Z}_M = \{x^* \in [a, b] : |f(x^*) - p^*(x^*)| = \max_{a \leqslant x \leqslant b} |f(x) - p^*(x)|\}.$$

We claim that the number of elements in $\mathscr{Z}_M$, denoted as $n(\mathscr{Z}_M)$, is at least $n+2$. Otherwise, if $n(\mathscr{Z}_M) \leqslant n+1$, we can show that there exists a function $p \in \mathscr{A}$ s.t.

$$[f(x) - p^*(x)]p(x) > 0, \quad \forall x \in \mathscr{Z}_M,$$

which contradicts our assumption by the conclusion in Theorem 1.

In fact, we can suppose that $\mathscr{Z}_M = \{\xi_1, \xi_2, \cdots, \xi_m\}$, where $m \leqslant n+1$ and $\xi_i \neq \xi_j$ for $i \neq j$. Let $\{\phi_i(x)\}_{i=1}^{n+1}$ be a basis of $\mathscr{A}$, and we assume that p could be expanded as the form

$$p(x) = \sum_{j=1}^{n+1} \alpha_j \phi_j(x), \quad \text{for some } \alpha_j,\ j = 1, 2, \cdots, n+1.$$

Let p satisfy $p(\xi_i) = \text{sign}[f(\xi_i) - p^*(\xi_i)]$, which leads to

$$p(\xi_i) = \sum_{j=1}^{n+1} \alpha_j \phi_j(\xi_i) = \text{sign}[f(\xi_i) - p^*(\xi_i)], \quad i = 1, 2, \cdots, m.$$

The linear system has non-zero solutions for α_j. Thus such a $p(x)$ does exist.

We claim again that $f(x) - p^*(x)$ changes sign at least $n+1$ times as x ranges over $\mathscr{Z}_M$. Otherwise, if $f(x) - p^*(x)$ changes sign at most n times, we suppose that its value at $\zeta_1, \zeta_2, \cdots, \zeta_{n+1} \in \mathscr{Z}_M$ has different signs, i.e.,

$$f(\zeta_i) - p^*(\zeta_i) = (-1)^i \lambda \|f - p^*\|_\infty, \quad i = 1, 2, \cdots, n+1,$$

where $\lambda = 1$ or -1. Then we can find $p \in \mathscr{A}$ such that

$$[f(\zeta_i) - p^*(\zeta_i)]p(\zeta_i) > 0, \quad i = 1, \cdots, n+1,$$

since any basis of $\mathscr{A}$ constitutes a Chebyshev set.

Hence we have shown that $f(x) - p^*(x)$ changes sign at least $n+1$ times as x ranges over $\mathscr{Z}_M$, which is the sufficient condition of this theorem.

Sufficiency. Suppose that conditions (11), (12) and (13) are satisfied, and then $p^* \in \mathscr{A}$ such that $f(x) - p^*(x)$ cahnges sign at least $n+1$ times as x ranges over $\mathscr{Z}_M$. We claim that there is no function $p \in \mathscr{A}$ such that

$$[f(\xi_i^*) - p^*(\xi_i^*)]p(\xi_i^*) > 0, \quad i = 0, 1, \cdots, n+1.$$

Otherwise, if such p does exist, then the sign of $p(\xi_i^*)$ is as same as the sign of $f(\xi_i^*) - p^*(\xi_i^*)$, i.e.,

$$\text{sign}(p(\xi_i^*)) = \text{sign}[f(\xi_i^*) - p^*(\xi_i^*)], \quad i = 0, 1, \cdots, n+1.$$

By the condition (13), we have

$$\text{sign}(p(\xi_{i+1}^*)) = -\text{sign}(p(\xi_i^*)), \quad i = 0, 1, \cdots, n.$$

Then $p(x)$ has at least $n+1$ zeros. By Haar conditions, we obtain $p \equiv 0$, which leads to a contradiction. Hence this claim is proved which implies p^* is the best minimax approximation by virtue of Theorem 1. □

Remark 3.4. *One important application of this theorem is to prove the minimum property of Chebyshev polynomials.*

We recall that the Chebyshev polynomial T_n is the polynomial of degree n that is defined on the interval $[-1, 1]$ by the equation

$$T_n(x) = \cos(n\theta), \quad x = \cos\theta, \quad 0 \leqslant \theta \leqslant \pi \tag{14}$$

The minimum property is sufficiently useful to be stated as a theorem.

Theorem 3. *Let the range of x be $[-1, 1]$, and let n be any positive integer. The polynomial $\left(\frac{1}{2}\right)^{n-1} T_n$ is the member of $\mathscr{P}_n$, whose ∞-norm is least, subject to the condition that the coefficient of x^n is equal to one.*

Proof. One way of identifying the required polynomial is to solve the problem

$$\begin{cases} \text{Find } \{c_i^* : i = 0, 1, \cdots, n-1\} \text{ s.t.} \\ \max\limits_{-1 \leqslant x \leqslant 1} \left| x^n + \sum\limits_{i=0}^{n-1} c_i^* x^i \right| \leqslant \max\limits_{-1 \leqslant x \leqslant 1} \left| x^n + \sum\limits_{i=0}^{n-1} c_i x^i \right|, \quad \forall c_i. \end{cases}$$

We see that this approach is equivalent to finding the best minimax approximation from $\mathbb{P}_{n-1}$ to the function x^n in the interval $[-1,1]$.

It follows from Theorem 2 that $(\frac{1}{2})^{n-1}T_n$ is the required polynomial, if the coefficient of x^n is one, and if there exist points $\{\xi_i : i = 0, 1 \cdots, n\}$ in $[-1,1]$ arranged in ascending order, such that the equations

$$T_n(\xi_i) = (-1)^{n-i}\|T_n\|_\infty, \quad i = 0, 1, \cdots, n, \tag{15}$$

hold. The recurrence relation of Tchebyshev polynomials implies that the coefficient of x^n is correct. Moreover, the definition (14) shows that equation (15) is satisfied if we let each ξ_i have the value $\cos[(n-i)\pi/n]$. The theorem is proved. □

Remark 3.5. *The main reason for letting $\mathscr{Z}$ be any closed subset of $\mathscr{C}[a,b]$ in the statement of Theorem 7.1, is that the exchange algorithm requires the case when $\mathscr{Z}$ contains just $(n+2)$ points. In descriptions of the exchange algorithm it is usual to call such a set of points a 'reference'. We use this term also, and we let $\{\xi_i; i = 0, 1, \ldots, n+1\}$ be the points of the reference. We assume that always these points are in ascending order*

$$a \leqslant \xi_0 < \xi_1 < \ldots < \xi_{n+1} \leqslant b \tag{7.23}$$

The following corollary of Theorem 7.1 is used on every iteration of the exchange algorithm.

Theorem 4. *Let $\mathscr{A}$ be an $(n+1)$-dimensional linear subspace of $\mathscr{C}[a,b]$ that satisfies the Haar condition, let $\{\xi_i; i = 0, 1, \ldots, n+1\}$ be a reference, and let f be any function in $\mathscr{C}[a,b]$. Then p^* is the function in $\mathscr{A}$ that minimizes the expression*

$$\max_{i=0,1,\ldots,n+1} |f(\xi_i) - p(\xi_i)|, \quad p \in \mathscr{A} \tag{16}$$

if and only if the equations

$$f(\xi_{i+1}) - p^*(\xi_{i+1}) = -[f(\xi_i) - p^*(\xi_i)], \quad i = 0, 1, \ldots, n \tag{17}$$

are satisfied.

Proof. Let $\mathscr{Z} = \{\xi_0, \xi_1, \cdots, \xi_{n+1}\}$, and $\mathscr{Z}_M \subset \mathscr{Z}$ be the set for which $|f(x) - p^*(x)|$ attains its maximum, i.e.,

$$\mathscr{Z}_M = \{x^* \in \mathscr{Z} : |f(x^*) - p^*(x^*)| = \max_{0 \leqslant i \leqslant n+1} |f(\xi_i) - p^*(\xi_i)|\}.$$

By Theorem 1, p^* is the best minimax approximation over the set $\mathscr{Z}$, i.e.,

$$\max_{0\leqslant i\leqslant n+1} |f(\xi_i) - p^*(\xi_i)| \leqslant \max_{0\leqslant i\leqslant n+1} |f(\xi_i) - p(\xi_i)|, \quad \forall p \in \mathscr{A},$$

if and only if there is no function $p \in \mathscr{A}$ such that

$$[f(x) - p^*(x)]p(x) > 0, \quad \forall x \in \mathscr{Z}_M.$$

Necessity. We claim that $\mathscr{Z}_M = \mathscr{Z}$. In fact, if it is not true, then the number of elements in $\mathscr{Z}_M$ is at most $n+1$, and then we can show there exists a function $p \in \mathscr{A}$ such that

$$[f(x) - p^*(x)]p(x) > 0, \quad \forall x \in \mathscr{Z}_M,$$

which leads to a contradiction. (See the proof of Theorem 2). Then we proved that $\mathscr{Z}_M = \mathscr{Z}$. Hence, we can show that $f(x) - p^*(x)$ changes sign at least $n+1$ times as x ranges over $\mathscr{Z}_M$. (See the proof of Theorem 2)

Sufficiency. We assume that (17) holds. Then it is obvious that $\mathscr{Z}_M = \mathscr{Z}$. We can show a proof by contradiction that there is no function $p \in \mathscr{A}$ such that

$$[f(x) - p^*(x)]p(x) > 0, \quad \forall x \in \mathscr{Z}_M,$$

holds. □

The function p^* that minimizes expression (16) may be calculated from the equations (17). It is usual to let h be the value of $[f(\xi_0) - p^*(\xi_0)]$, and to choose a basis of $\mathscr{A}, \{\phi_j; j = 0, 1, \ldots, n\}$ say. It follows that the coefficients of the function

$$p^*(x) = \sum_{j=0}^{n} \lambda_j \phi_j(x), \quad a \leqslant x \leqslant b \tag{7.26}$$

satisfy the equations

$$f(\xi_i) - \sum_{j=0}^{n} \lambda_j \phi_j(\xi_i) = (-1)^i h, \quad i = 0, 1, \ldots, n+1 \tag{7.27}$$

which is a linear system in the unknowns $\{\lambda_j; j = 0, 1, \ldots, n\}$ and h. Because Theorem 4 shows that these equations have a solution for all functions f in $\mathscr{C}[a, b]$, the matrix of the system is non-singular. Hence only one element of $\mathscr{A}$ minimizes expression (16). A more general and more useful method of proving uniqueness is given in the next section.

4 Uniqueness and bounds on the minimax error

Suppose that the conditions of Theorem 2 hold, that p^* and q^* are both best minimax approximations from $\mathscr{A}$ to f, and that conditions (11), (12) and (13) are satisfied. We let r^* be the function $(q^* - p^*)$, and we consider the numbers

$$r^*(\xi_i^*) = [f(\xi_i^*) - p^*(\xi_i^*)] - [f(\xi_i^*) - q^*(\xi_i^*)], \quad i = 0, 1, \ldots, n+1 \tag{18}$$

Because $\|f - q^*\|_\infty$ and $\|f - p^*\|_\infty$ are equal, it follows from equation (12) that either $r^*(\xi_i^*)$ is zero, or its sign is the same as the sign of $[f(\xi_i^*) - p^*(\xi_i^*)]$. Hence equation (13) provides information about the signs of the terms of the sequence $\{r^*(\xi_i^*); i = 0, 1, \ldots, n+1\}$.

It can be deduced from this information that r^* is identically zero. (Since $r^* \in \mathscr{A}$ and it has at least $n+1$ zeros, by Haar condition, which leads to it identically zero).

Hence the best minimax approximation from $\mathscr{A}$ to f is unique. The method of proving that r^* is identically zero is a general one that has several applications. Therefore it is stated in the following theorem.

Theorem 5. *Let r be a function in $\mathscr{C}[a, b]$, and let $\{\xi_i : i = 0, 1, \ldots, n+1\}$ be a reference, such that the conditions*

$$(-1)^i r(\xi_i) \geqslant 0, \quad i = 0, 1, \ldots, n+1 \tag{19}$$

are satisfied. Then r has at least $(n+1)$ zeros in $[a, b]$, provided that any double zero is counted twice, where a double zero is a zero that is strictly inside $[a, b]$, at which r does not change sign.

Proof. Let $\mathscr{I}$ and $\mathscr{J}$ be the sets

$$\left.\begin{array}{ll} \mathscr{I} = \{i : r(\xi_i) \neq 0, & i = 0, 1, \ldots, n+1\} \\ \mathscr{J} = \{j : r(\xi_i) = 0, & j = 0, 1, \ldots, n+1\} \end{array}\right\}, \tag{20}$$

and let $n(\mathscr{I})$ and $n(\mathscr{F})$ be the number of elements in each set. The theorem is trivial if $n(\mathscr{F})$ is zero or one. Otherwise we consider the number of zeros in the interval $[\xi_k, \xi_l]$, where k and l are both in $\mathscr{I}$, and where no other element of $\mathscr{I}$ is in the range $[k, l]$. Condition (19) implies that the numbers $r(\xi_k)$ and $r(\xi_l)$ have the same sign if $(l - k)$ is even, and they have opposite signs if $(l - k)$ is odd. Hence the number of zeros of r in the interval $[\xi_k, \xi_l]$ is at least one more than the number of points of the set $\{\xi_j; j \in \mathscr{J}\}$ that are in this interval, provided that any double zero is

counted twice. Because the number of pairs $[\xi_k, \xi_l]$ that have this property is $[n(\mathscr{I}) - 1]$, it follows that the total number of zeros of r in $[a, b]$ is at least $[n(\mathscr{I}) + n(\mathscr{J}) - 1]$, which is the required result. $\square$ $\square$

Hence we obtain the uniqueness theorem for best approximation in the ∞-norm.

Theorem 6. *Let $\mathscr{A}$ be a linear subspace of $\mathscr{C}[a, b]$ that satisfies the Haar condition. Then, for any f in $\mathscr{C}[a, b]$, there is just one best minimax approximation from $\mathscr{A}$ to f.*

Proof. Suppose that p^* and q^* are both best approximations. Let $r^*(x) = (p^*(x) - q^*(x))$. Then by Theorem 2, there exist $(n + 2)$ points $\{\xi_i^* : i = 0, 1, \cdots, n + 1\}$ such that (11), (12) and (13) are satisfied. Then for $i = 0, 1, \cdots, n + 1$,

$$|f(\xi_i^*) - q^*(\xi_i^*)| \leqslant \|f - q^*\|_\infty = \|f - p^*\|_\infty = |f(\xi_i^*) - p^*(\xi_i^*)|.$$

By expression (18), $r^*(\xi_i^*)$ is either zero or

$$\text{sign}(r^*(\xi_i^*)) = \text{sign}(f(\xi_i^*) - p^*(\xi_i^*)).$$

Therefore,

$$\text{sign}(r^*(\xi_{i+1}^*)) = -\text{sign}(r^*(\xi_i^*)), \quad i = 0, 1, \cdots, n,$$

where we define $\text{sign}(0) = 1$. Then

$$(-1)^i \lambda r^*(\xi_i^*) \geqslant 0, \quad i = 0, 1, \cdots, n + 1,$$

for which $\lambda = 1$ or -1 is a constant. By Theorem 5, $r^*(x)$ has at least $n + 1$ zeros in $[a, b]$ and $r^*(x) \in \mathscr{A}$. It follows from property (1) of Haar condition, that $r^*(x)$ is identically zero. $\square$

Remark 4.1. *Another interesting property of the Haar condition, which is the subject of Exercise 7.9, is that, if $\mathscr{A}$ is any finite-dimensional linear subspace of $\mathscr{C}[a, b]$ that does not satisfy the Haar condition, then there are functions f in $\mathscr{C}[a, b]$ that have several best approximations in $\mathscr{A}$.*

Theorem 5 is also useful for obtaining lower bounds on the least value of expression (1). Suppose that an iterative method for calculating a best approximation produces a trial approximation p^*, and that the conditions (11), (12) and (13) are almost satisfied. Then we usually have available a reference $\{\xi_i : i = 0, 1, \ldots, n + 1\}$, such that the signs of the terms

$$\{f(\xi_i) - p^*(\xi_i) : i = 0, 1, \ldots, n + 1\}$$

alternate. In this case the following theorem applies.

Theorem 7. *Let the conditions of Theorem 2 hold, let p^* be any element of $\mathscr{A}$, and let $\{\xi_i; i = 0, 1, \ldots, n+1\}$ be a reference, such that the condition*

$$\operatorname{sign}\left[f\left(\xi_{i+1}\right) - p^*\left(\xi_{i+1}\right)\right] = -\operatorname{sign}\left[f\left(\xi_i\right) - p^*\left(\xi_i\right)\right], \quad i = 0, 1, \ldots, n \tag{21}$$

is satisfied. Then the inequalities

$$\begin{aligned} \min_{i=0,1,\ldots,n+1} \left|f\left(\xi_i\right) - p^*\left(\xi_i\right)\right| &\leqslant \min_{p\in\mathscr{A}} \max_{i=0,1,\ldots,n+1} \left|f\left(\xi_i\right) - p\left(\xi_i\right)\right| \\ &\leqslant \min_{p\in\mathscr{A}} \|f - p\|_\infty \\ &\leqslant \|f - p^*\|_\infty \end{aligned} \tag{22}$$

are obtained. Moreover, the first inequality is strict unless all the numbers $\{|f\left(\xi_i\right) - p^*\left(\xi_i\right)| : i = 0, 1, \ldots, n+1\}$ are equal.

Proof. The third inequality of expression (22) holds because p^* is in $\mathscr{A}$, and the second one holds because the reference is a subset of $[a, b]$. In the next we prove the first inequality, i.e.,

$$\min_{0\leqslant i\leqslant n+1} |f(\xi_i) - p^*(\xi_i)| \leqslant \min_{p\in\mathscr{A}} \max_{0\leqslant i\leqslant n+1} |f(\xi_i) - p(\xi_i)|. \tag{23}$$

By contradiction, we suppose that (23) is not hold, i.e.,

$$\min_{0\leqslant i\leqslant n+1} |f(\xi_i) - p^*(\xi_i)| > \min_{p\in\mathscr{A}} \max_{0\leqslant i\leqslant n+1} |f(\xi_i) - p(\xi_i)|.$$

Since $\mathscr{A}$ is finite dimensional and then it is a closed subspace. There exists $q^*(x) \in \mathscr{A}$ such that

$$\min_{0\leqslant i\leqslant n+1} |f(\xi_i) - p^*(\xi_i)| > \max_{0\leqslant i\leqslant n+1} |f(\xi_i) - q^*(\xi_i)|.$$

Then we have

$$|f(\xi_i) - p^*(\xi_i)| > |f(\xi_i) - q^*(\xi_i)|, \quad i = 0, 1, \cdots, n+1.$$

We set $r^* = q^* - p^*$ as defined in (18), then it is obvious that $\operatorname{sign}(r^*(\xi_i)) = \operatorname{sign}(f(\xi_i) - p^*(\xi_i))$ and by (21), that r^* has at least $n+1$ zeros in $[a, b]$. It follows that $r^* \equiv 0$ from $r^* \in \mathscr{A}$ and condition (1) of Haar condition. Therefore $q^* = p^*$, which leads to

$$\min_{0\leqslant i\leqslant n+1} |f(\xi_i) - p^*(\xi_i)| > \max_{0\leqslant i\leqslant n+1} |f(\xi_i) - p^*(\xi_i)|,$$

which is a contradiction.

Finally, we prove the proposition: the equal sign of (23) holds if and only if all the numbers $\{|f(\xi_i) - p^*(\xi_i)| : i = 0, 1, \ldots, n+1\}$ are equal.

Necessity

We suppose that

$$\min_{0 \leqslant i \leqslant n+1} |f(\xi_i) - p^*(\xi_i)| = \min_{p \in \mathscr{A}} \max_{0 \leqslant i \leqslant n+1} |f(\xi_i) - p(\xi_i)|.$$

Then there exists $q^* \in \mathscr{A}$ such that

$$\min_{0 \leqslant i \leqslant n+1} |f(\xi_i) - p^*(\xi_i)| = \max_{0 \leqslant i \leqslant n+1} |f(\xi_i) - q^*(\xi_i)|.$$

We can prove that $q^* = p^*$ by contradiction. Then

$$\min_{0 \leqslant i \leqslant n+1} |f(\xi_i) - p^*(\xi_i)| = \max_{0 \leqslant i \leqslant n+1} |f(\xi_i) - p^*(\xi_i)|,$$

which leads to $|f(\xi_i) - p^*(\xi_i)|$ are the same.

Sufficiency

By Theorem 4.

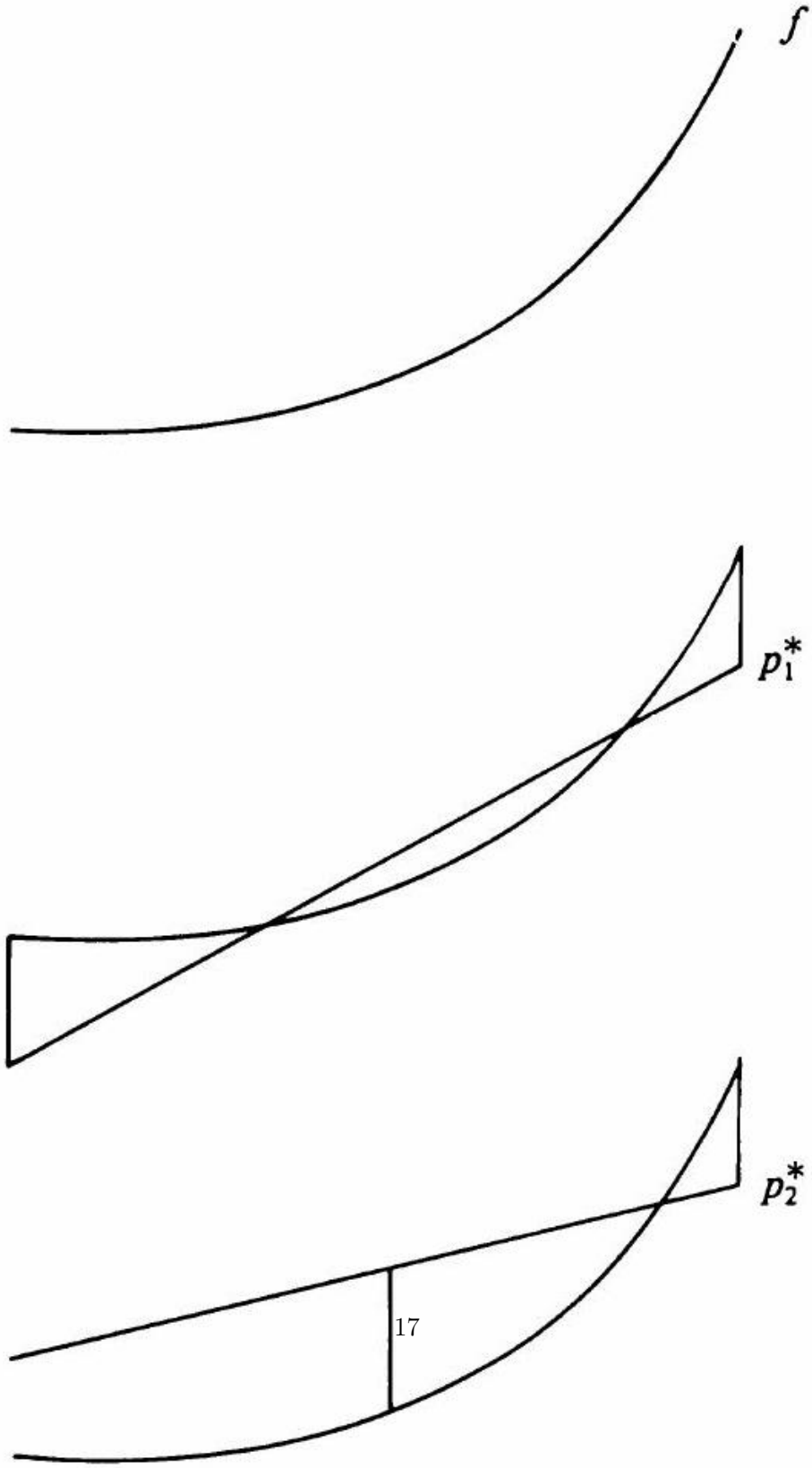
□

It is useful to note that, if p^* is the best minimax approximation from $\mathscr{A}$ to f, and if the reference in the statement of the last theorem is the set of points $\{\xi_i^*; i = 0, 1, \ldots, n+1\}$ that occurs in conditions (11), (12) and (13), then all the inequalities of expression (22) are satisfied as equations.

5 Appendix

Notation Interpretation:

- $\mathscr{A}$ Linear space of $\mathscr{C}[a, b]$


f
p_1^*
p_2^*
17